La pseudonimizzazione come nuova garanzia di accountability nell'universo della profilazione e dei Big Data

Date : 9 gennaio 2018

La possibilità di generare e acquisire informazioni sugli individui nella società dell'informazione ha messo a dura prova la regolamentazione della protezione dei dati personali.

Le minacce alla privacy sono sempre maggiori e incombenti a causa del progresso tecnologico che può da un lato compromettere la sicurezza dei dati che vengono conservati, dall'altro consentire ai titolari del trattamento di aggregare informazioni acquisite su uno stesso individuo senza che il soggetto cui esse appartengono sia consapevole di poter essere identificato o di rendersi identificabile per il fatto di averle conferite.

Proprio per queste ragioni, viste le preoccupazioni derivanti da fenomeni quali le analisi, spesso predittive, dei comportamenti online degli utenti, a partire dai Big Data raccolti anche tramite l'Internet of Things e il conseguente ampliamento delle tradizionali banche dati, si è giunti a parlare in modo frequente di "anonimato" e di "tecniche di anonimizzazione" dei dati. Queste ultime, infatti, permetterebbero, data la loro irreversibilità, di rendere l'interessato non più identificato o identificabile.

Tuttavia, grazie al Regolamento Europeo in materia di protezione dei dati personali (di seguito, "GDPR"), è stata introdotta una nuova soluzione che favorisce la tutela dell'individuo e dei suoi dati personali: si tratta della pseudonimizzazione, cioè «*il trattamento dei dati personali in modo tale che i dati personali non possano più essere attribuiti a un interessato specifico senza l'utilizzo di informazioni aggiuntive, a condizione che tali informazioni aggiuntive siano conservate separatamente e soggette a misure tecniche e organizzative intese a garantire che tali dati personali non siano attribuiti a una persona fisica identificata o identificabile*» (art. 4.5), GDPR). In sostanza, la pseudonimizzazione implica tre elementi:

- l'assenza di identificabilità diretta del soggetto interessato («*trattamento dei dati personali in modo tale che i dati non possano essere più attribuiti ad un interessato specifico senza l'utilizzo di informazioni aggiuntive*»);

- l'adozione di misure di sicurezza ulteriori da aggiungere alla pseudonimizzazione («*a*

condizione che tali informazioni aggiuntive siano conservate separatamente»);

- l'incorporazione della pseudonimizzazione nella privacy-by-design («*e soggette a misure tecniche e organizzative intese a garantire che tali dati personali non siano attribuiti a una persona fisica identificata o identificabile*»).

Il Considerando 28 del GDPR sottolinea, infatti che «*L'applicazione della pseudonimizzazione ai dati personali può ridurre i rischi per gli interessati e aiutare i titolari del trattamento e i responsabili del trattamento a rispettare i loro obblighi di protezione dei dati*». Tale utilità è confermata dall'art. 32, relativo alla Sicurezza del trattamento, il quale inserisce tra le «*misure tecniche e organizzative adeguate per garantire un livello di sicurezza adeguato al rischio*» proprio la pseudonimizzazione. Chiaramente, e come precisato dal citato articolo, ci sono anche altre misure di sicurezza adottabili. Tuttavia il valore della pseudonimizzazione sta nel non essere solo una tecnica da combinarsi con ulteriori misure di sicurezza, ma nell'avere un efficacia che consente di considerarla già di per sé una misura adottata a tutela dei dati personali dei soggetti interessati, per diminuirne i rischi di identificazione diretta.

Dunque, tornando ai Big Data, la pseudonimizzazione consente di raccogliere dati diversi ma relativi allo stesso soggetto, senza che di esso si conosca l'identità in modo diretto. Così, anche se il soggetto rimane identificabile, devono comunque sussistere motivi legittimi per effettuare la reidentificazione in quanto i dati personali devono essere «*raccolti per finalità determinate, esplicite e legittime, e successivamente trattati in modo che non sia incompatibile con tali finalità*» (Art. 5.1.b), GDPR).

Volendo soffermarsi sulle operazioni di reidentificazione, giacché pseudonimizzare i dati implica che si possano effettuare simili operazioni, esse dovranno essere ricomprese tra le finalità del trattamento comunicate al soggetto al momento della raccolta dei dati; al tempo stesso, però, sarà la pseudonimizzazione stessa a ridurre al minimo il trattamento dei dati personali conformemente ai principi di proporzionalità e di necessità, il che le rende sia strumento di efficiente trattamento di dati sia garanzia della sua bassa invasività.

Ma quali sono, concretamente, le garanzie offerte dalla pseudonimizzazione quando si parla di Big Data e profilazione? In cosa differisce dall'anonimizzazione e perché ha assunto una simile rilevanza nelle disposizioni europee?

Naturalmente, sia la pseudonimizzazione che l'anonimizzazione vengono poste a tutela del singolo individuo, inteso come "soggetto identificabile", al fine di garantirgli protezione rispetto alle attività di profilazione mirate che comportano l'identificazione del soggetto (*single out*). È noto infatti che *ex* art. 22.1, GDPR «*L'interessato ha il diritto di non essere sottoposto a una decisione basata unicamente sul trattamento automatizzato, compresa la profilazione, che produca effetti giuridici che lo riguardano o che incida in modo analogo significativamente sulla sua persona*».

È innegabile, tuttavia, che il cd. *data mining*, avvalendosi dei Big Data e delle tecniche di profilazione, ha il merito di combinare tra loro le informazioni in modo da formulare inferenze su eventi altrimenti difficilmente prevedibili, sia che si tratti di calamità naturali, epidemie o

comportamenti individuali.

Eppure, la *ratio* sottesa alla protezione dei dati deve tenere conto non solo del cd. *single out* (individuazione del comportamento individuale), ma anche del possibile inserimento del singolo all'interno di un *cluster* di individui, cioè della sua collocazione in gruppi che costituiscono categorie omogenee. Le possibilità offerte dalle operazioni di *data mining*, infatti, non riguardano solo la profilazione di un singolo individuo, ma anche il trattamento di dati funzionale all'individuazione di target/gruppi entro cui ricomprenderlo.

In tale contesto, l'aggregazione dei dati permessa dall'anonimizzazione permette di attenuare il *single out* nell'ambito di quello che potrebbe definirsi "*cluster bombing*". Inserendo il singolo individuo all'interno di un *cluster* anonimizzato si impedisce, in effetti, l'identificazione del soggetto, permettendo comunque le operazioni di profilazione e di individuazione di gruppi/target estesi. Mediante l'utilizzo dell'aggregazione non si ha più un impatto concentrato sul singolo componente del cluster (identificazione), poiché l'obiettivo amplia la sua portata, estendendosi all'intero gruppo.

Date queste premesse, si può meglio comprendere il valore della pseudonimizzazione e come essa si inserisce tra reidentificazione, *clustering* e responsabilità del titolare del trattamento.

Quando i dati degli individui vengono anonimizzati inserendo i soggetti all'interno di un *cluster*, il grado di incontrollabilità degli impatti della profilazione diretta a target di individui aumenta esponenzialmente. Per utilizzare una metafora, è come avere un cecchino che, durante il Carnevale di Venezia, non conoscendo l'identità del suo bersaglio (target) poiché tutti coloro che affollano la piazza indossano la stessa maschera, inizia a sparare all'intera folla, nascosto sui tetti. Questo comporterebbe non solo il coinvolgimento di più vittime rispetto a quella cui effettivamente sarebbe diretto il colpo di pistola, ma anche l'impossibilità di risalire all'identità dell'assassino giacché egli non possiede un movente per ciascuna delle persone colpite e le indagini finirebbero per aprire piste inconcludenti. Fuor di metafora, nel caso dell'anonimizzazione finalizzata all'aggregazione dei dati di un *cluster*, il rischio è quello di non poter ricostruire i processi che hanno condotto all'individuazione del target e, quindi, di non poter risalire a colui che detiene le responsabilità rispetto al trattamento dei dati personali.

Si giunge, così, ad apprezzare e comprendere a pieno il valore della pseudonimizzazione nella società dei Big Data. Con la pseudonimizzazione, infatti, il rischio di non poter risalire a "colui che detiene le responsabilità rispetto al trattamento dei dati personali" viene scongiurato poiché uno o più soggetti assumono la funzione di "custodi dei dati" – come peraltro avviene già nella biometria nel settore bancario. Si ricorda, infatti, che la pseudonimizzazione si basa anche sulla misura per cui «*le informazioni aggiuntive per l'attribuzione dei dati personali a un interessato specifico siano conservate separatamente*» (*ex* Considerando 29, GDPR). Ma proprio grazie al fatto che la pseudonimizzazione garantisce la ricostruibilità dei processi di mascheramento dell'identità, permettendo la reidentificazione, essa assicura l'accountability, la responsabilizzazione *ex* art. 5.2, GDPR del titolare del trattamento. La reidentificazione, infatti, non è il punto debole delle tecniche di pseudonimizzazione: è una forma di tutela per il soggetto, sia inteso come singolo individuo che come possibile membro di un *cluster*.

Concludendo la presente riflessione, sono forse più chiari e apprezzabili i motivi per cui la pseudonimizzazione ha assunto enorme rilevanza all'interno del GDPR. «*L'applicazione della pseudonimizzazione ai dati personali può ridurre i rischi per gli interessati e aiutare i titolari del trattamento e i responsabili del trattamento a rispettare i loro obblighi di protezione dei dati*», ricorda il Considerando 28 del GDPR. La responsabilità del titolare del trattamento nelle operazioni di profilazione, infatti, viene garantita dalla reversibilità stessa del processo di pseudonimizzazione. Pertanto, ciò che può apparire come un rischio potenziale per l'individuo (possibilità di reidentificazione), diventa invece garanzia di tutela rispetto non solo al *single out* ma anche nell'ambito della clusterizzazione, assicurando che l'utilizzo dei dati da parte del titolare avvenga in maniera conforme alla normativa.

Bibliografia:

1. Regolamento (UE) 2016/679, in http://eur-lex.europa.eu/legal-content/IT/TXT/?uri=CELEX%3A32016R0679
2. Bolognini, C. Bistolfi, Pseudonymization and impacts of Big (personal/anonymous) Data processing in the transition from the Directive 95/46/EC to the new EU General Data Protection Regulation, Computer Law & Security review, 2016

A cura di: **Camilla Bistolfi**